

N. Leigh Anderson¹
Ricardo Esquer-Blasco²
Jean-Paul Hofmann¹
Lydie Meheus²
Jos Raymackers²
Sandra Steiner³
Frank Witzmann⁴
Norman G. Anderson¹

¹Large Scale Biology Corporation,
Rockville, MD

²Innogenetics NV, Ghent

³Sandoz Pharma Ltd, Drug Safety
Assessment, Toxicology, Basel

⁴Molecular Anatomy Laboratory,
Indiana University Purdue
University Columbus, Columbus, IN

An updated two-dimensional gel database of rat liver proteins useful in gene regulation and drug effect studies

We have improved upon the reference two-dimensional (2-D) electrophoretic map of rat liver proteins originally published in 1991 (N. L. Anderson *et al.*, *Electrophoresis* 1991, 12, 907-930). A total of 53 proteins (102 spots) are now identified, many by microsequencing. In most cases, spots cut from wet, Coomassie Blue stained 2-D gels were submitted to internal tryptic digestion [2], and individual peptides, separated by high-performance liquid chromatography (HPLC), were sequenced using a Perkin-Elmer 477A sequencer. Additional spots were identified using specific antibodies.

Figure 1 shows the current annotated 2-D map of F344 rat liver, analyzed using the Iso-DALT system (20 × 25 cm gels) and BDH 4-8 carrier ampholytes. Both the map itself and the master spot number system remain the same as shown in the original publication. Table 1 lists the important features of each identification shown, including the gel position, *pI*, and *M_r*, for the most abundant or most basic form of each protein. Using this extended base of identified spots, a series of four improved calibration functions has been derived for the *pI* and SDS-*M_r* axes (the first two of which are shown in Fig. 2A and B). Both forward and reverse functions are derived, so that one can compute the physical properties of a spot with a given gel location, or inversely compute the gel position expected for a protein having given physical properties:

$$Y_{\text{RATLIVER}} = f_{M_r-\text{RATLIVER}}(M_{r,\text{SEQUENCE-DERIVED}}) \quad (1)$$

$$X_{\text{RATLIVER}} = f_{pI-\text{RATLIVER}}(pI_{\text{SEQUENCE-DERIVED}}) \quad (2)$$

$$M_{r,\text{GEL-DERIVED}} = f_{\text{RATLIVER } Y-M_r}(Y_{\text{RATLIVER}}) \quad (3)$$

$$pI_{\text{GEL-DERIVED}} = f_{\text{RATLIVER } X-pI}(X_{\text{RATLIVER}}) \quad (4)$$

A spreadsheet program (in Microsoft Excel) was developed to facilitate flexible computation of *pI*'s from amino acid sequence data, and the results were entered into a relational database (Microsoft Access). A table of spot positions and sequence-derived *pI*'s and *M_r*'s was fitted with a large series of analytic equations using Tablecurve (Jandel Scientific), and the four conversion Eqs. (1)-(4), relating computed *pI* and gel *X* coordinate, or computed molecular weight and gel *Y* coordinate, were selected, based on criteria of simplicity, goodness of fit and favorable asymptotic behavior. Table 2 lists the equations and coefficients. Application of Eqs. (3) and (4) to a spot's *X* and *Y* coordinates, given in [1], produce improved *M_r* estimates, and allow computation of *pI*

directly in *pH* units, instead of in terms of positions relative to creatine phosphokinase (CPK) charge standards. The inverse Eqs. (1) and (2) were used to compute the gel positions of a series of *pI* and *M_r* tick marks. These tick marks were plotted with SigmaPlot (Jandel), together with fiducial marks locating several prominent spots, and the resulting graphic was aligned over the synthetic gel image (computed by Kepler from the master gel pattern) using Freelance (Lotus Development). Maps were printed as Postscript output from Freelance, either in black and white (as shown here) or in color, where label color indicates subcellular location (available from the first author upon request). We have also used the rat liver 2-D pattern as presented here to calibrate the patterns of other samples. Using mixtures of rat liver and mouse liver samples, for example, we made composite 2-D patterns that allow use of the rat pattern to standardize both axes of the mouse pattern. This was accomplished by deriving transformations relating the rat and mouse *X*, and separately the rat and mouse *Y*, axes (Table 2, lower half; Fig. 2C and D) based on a series of spots that coelectrophorese in these closely related species. These functions were then applied to derive equations relating the mouse liver *X* and *Y* to *pI* and SDS-*M_r* (Eqs. 5 and 6 below). The resulting standardized 2-D pattern for B6C3F1 mouse liver is shown in Fig. 3.

$$M_{r,\text{MOUSELIVER}} = f_{\text{RATLIVER } Y-M_r}(f_{\text{MOUSELIVER } Y-\text{RATLIVER } Y}(Y_{\text{MOUSELIVER}})) \quad (5)$$

$$pI_{\text{MOUSELIVER}} = f_{\text{RATLIVER } X-pI}(f_{\text{MOUSELIVER } X-\text{RATLIVER } X}(X_{\text{MOUSELIVER}})) \quad (6)$$

A slightly more complex approach can be used to standardize samples that have few or no spots co-electrophoresing with rat liver proteins. In this case, a 2-D gel is prepared with a mixture of the two samples, and four functions (forward and backward, each for *X* and *Y*) are derived relating each sample's own master pattern to the composite. The required functions are then applied in a nested fashion to yield the desired result (using rat plasma as an example):

$$M_{r,\text{RATPLASMA}} = f_{\text{RATLIVER } Y-M_r}(f_{\text{RATPLASMA+LIVER } Y-\text{RATLIVER } Y}(f_{\text{RATPLASMA } Y-\text{RATPLASMA+LIVER } Y}(Y_{\text{RATPLASMA}})))) \quad (7)$$

Correspondence: Dr. Leigh Anderson, Large Scale Biology Corporation, 9620 Medical Center Drive, Rockville, MD 20850-3338 USA (Tel: +301-424-5989; Fax: +301-762-4892; email: leigh@lsbc.com)

Keywords: Two-dimensional polyacrylamide gel electrophoresis / Liver / Map / Identification / Calibration

F344 RAT LIVER 2-D PROTEIN PATTERN

v1.6 (F344MST3.ms29) 28-Apr-1995 © by Large Scale Biology Corporation,
9620 Medical Center Drive, Rockville, MD 20850 USA 301/424-5989
MW and computed pI scales derived from 616 known proteins

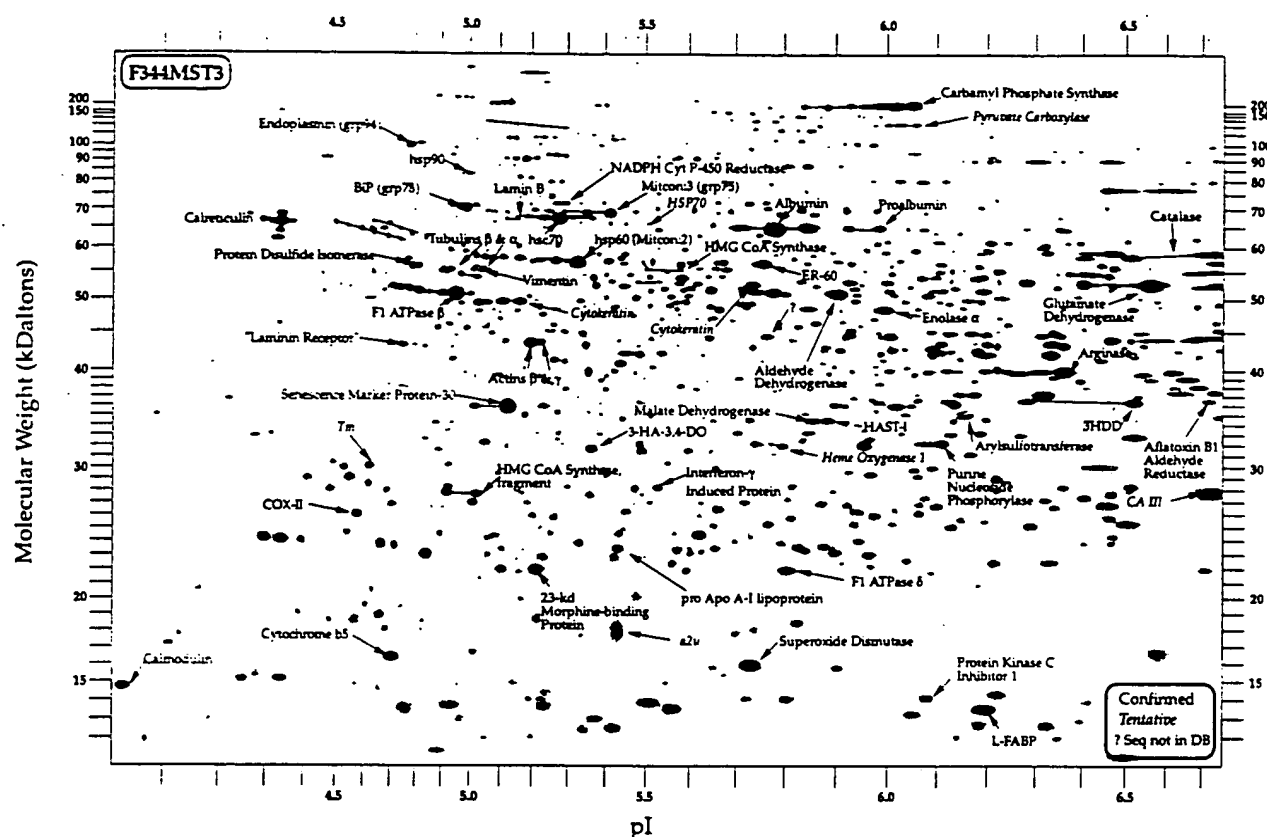


Figure 1. Master 2-D gel pattern of Fischer 344 rat liver proteins, annotated with 53 protein identifications and computed pI and M_r axes. Tentative identifications are in italic type.

Table 1. Proteins identified in the 2-D pattern of F344 rat liver

MSN ^{a)}	Protein ID ^{b)}	Protein name	Identification comments	Gel X ^{c)}	Experimental pI ^{d)}	Gel Y ^{c)}	Experimental M_r ^{d)}
126	HADO-HUMAN ^{a)}	3-HA-3,4-DO: 3-hydroxy-anthranilate-3,4-dioxygenase	Internal sequence	871.95	5.36	921.35	30 207
137, 159, 288, 258	DIDH_RAT	3HDD: 3-hydroxysteroid dihydrodiol reductase	Ab (T.M. Penning) and pure protein	1857.52	6.51	822.52	34 406
173	MUP_RAT	$\alpha_2\mu$ globulin	Presence in liver microsome lumen, abundance in kidney, pI, M_r	919.16	5.43	1313.81	19 549
38	ACTB_HUMAN	Actin β	Analogy with other mammalian patterns (e.g. human) through coelectrophoresis	763.40	5.19	693.64	41 586
68	ACTG_HUMAN	Actin γ	Analogy with other mammalian patterns (e.g. human) through coelectrophoresis	779.42	5.21	692.26	41 677
693	AFAR_RAT	Aflatoxin B1 aldehyde reductase	Internal sequence	1993.32	6.72	818.60	34 593
28, 21, 33	ALBU_RAT	Albumin	Coelectrophoresis with principal plasma protein	1262.81	5.86	445.64	66 354
43	DHAM_RAT	Aldehyde dehydrogenase	N-Terminal sequence and AAA	1317.72	5.91	589.03	49 602
96	ARGI_RAT	Arginase	Internal sequence	1730.72	6.34	756.02	37 819
117	SUAR_RAT	Arylsulfotransferase	Internal sequence	1547.96	6.14	849.08	33 186
1163, 1161, 1162, 20	GR78_RAT	BIP (GRP-78)	Ab (F. Witzmann)	665.33	5.01	397.39	74 564
185	CAH3_RAT	CA-III	Uncertain; by comparison with mouse	1996.60	6.72	1017.02	26 887
123	CALM_HUMAN	Calmodulin	Analogy with human cellular patterns through coelectrophoresis	23.05	4.03	1433.25	17 419
3, 201, 48, 39, 22, 24	CRTC_RAT	Calreticulin	Ab (Lance Pohl)	310.59	4.34	433.80	68 206

Table 1. continued

MSN ^{a)}	Protein IDb)	Protein name	Identification comments	Gel X ^{c)}	Experimental pI ^{d)}	Gel Y ^{c)}	Experimental M _r ^{d)}
1184, 1186, 114, 174, 118 5, 167, 157 54, 61 136	CPSM_RAT CATA_RAT COX2_RAT	Carbamyl phosphate synthase Catalase COX-II	2-D of pure protein; confirmed by N-terminal sequence and AAA Internal sequence Ab (J. W. Taanman), confirmed by internal sequence	1453.56 2000.81 452.57	6.05 6.73 4.61	181.64 499.64 1062.67	160 640 58 968 25 504
87	CYB5_RAT	Cytochrome B5	2-D of pure protein; Ab; confirmed by AAA	515.68	4.73	1370.55	18 493
41 29	CK-RAT ^{e)} CK-RAT ^{e)}	Cytokeratin Cytokeratin	Location in cytoskeletal fraction Location in cytoskeletal fraction	1165.12 743.11	5.75 5.15	569.09 605.23	51 448 48 187
5, 11 60	ENPL-RAT ^{e)} ENOA_RAT	Endoplasmic Enolase A	Ab (F. Witzmann) Internal sequence and AAA	567.73 1399.78	4.83 6.00	263.37 623.54	112 194 46 674
27 17	ER60_RAT ATPB_RAT	ER-60 F1 ATPase β	N-Terminal sequence (R. M. Van Frank) N-Terminal sequence and AAA	1184.20 629.06	5.77 4.95	523.51 588.83	56.169 49 620
196 79	ATP7_RAT F16P_RAT	F1 ATPase δ Fructose-1,6-bis-phosphatase	Internal sequence Uncertain; by comparison with ID in Garrison and Wager (JBC 257:13135-13143)	1227.24 924.54	5.82 5.44	1184.65 737.77	22 310 38 858
62, 78 125	DHE3_RAT HAST-RAT ^{e)}	Glutamate dehydrogenase HAST-I: N-hydroxyaryl-amine sulfotransferase	N-Terminal sequence and internal sequence Internal sequence	1887.39 1297.94	6.55 5.89	566.92 861.55	51 655 32 638
307	HO1_RAT	Heme oxygenase 1	Uncertain; available data from internal sequence	1219.39	5.81	915.71	30 423
413, 1250, 933	HMCS_RAT	HMG CoA synthase, cytosolic	Ab (J. Gersmshausen)	1033.48	5.59	538.13	54 571
133, 144, 235	HMCS_RAT	HMG CoA synthase, mitochondrial (frag)	Ab (J. Gersmshausen), N-terminal sequence (Steiner/Lottspeich)	666.40	5.02	1019.42	26 811
8, 23, 1307	HS7C_RAT	HSC-70	Positional homology (with human, etc.) through coelectrophoresis	811.87	5.27	425.76	69 521
15, 25, 110	P60_RAT	HSP-60	Ab (F. Witzman); confirmed by N-terminal sequence and AAA	845.09	5.32	520.03	56 561
971 1216, 1215, 90 256	HS70-RAT ^{e)} HS90-RAT ^{e)} ING1-HUMAN	HSP-70 HSP-90 Interferon- γ induced protein	Ab (F. Witzman) Ab (F. Witzman) Internal sequence	976.11 659.86 993.85	5.51 5.00 5.54	437.14 329 1006.04	67 674 90 107 27 237
415, 734	LAMB-RAT ^{e)}	Lamin B	Positional homology with human through coelectrophoresis, nuclear location	737.10	5.14	425.19	69 615
80 227	LAMR-RAT ^{e)} FABL_RAT	"Laminin receptor" L-FABP (liver fatty acid binding protein)	Internal sequence Ab (N. M. Bass)	534.02 1586.09	4.77 6.18	697.62 1483.43	41 327 16 622
134	MDHC_MOUSE E	Malate dehydrogenase	Internal sequence	1270.85	5.86	861.96	32 620
18, 35, 226	GR75-RAT ^{e)}	Mitcon-3; grp75	Positional homology with human through coelectrophoresis	905.67	5.41	413.67	71 589
175, 251 1168, 1170, 1171	NCPR_RAT PDI_RAT	NADPH P450 reductase PDI: Protein disulfide isomerase	2-D of pure protein N-Terminal sequence (R. M. van Frank), Ab	824.69 564.30	5.29 4.83	393.21 528.47	75 366 55 618
47, 93	ALBU_RAT	Pro-Albumin	Microsomal lumen location, pI, M _r relative to albumin	1391.03	5.99	446.68	66 195
236 320	APA1_RAT IPK1_BOVIN	Pro-APO A-I lipoprotein Protein kinase C inhibitor 1	Coelectrophoresis with plasma protein Internal sequence; homology with bovine protein	920.41 1480.01	5.43 6.08	1137.51 1458.81	23 467 17 007
152	PNPH_MOUSE	Purine nucleoside phosphorylase	Internal sequence	1507.19	6.10	911.16	30 599
1179, 1180, 1181, 1182, 1183	PYVC-RAT ^{e)}	Pyruvate carboxylase	Tentative; 2-D of pure protein (J. G. Henslee, JBC, 1979); reported in <i>Biochim. Biophys. Acta</i> 1022, 115-125	1485.10	6.08	223.52	131 589
55, 103	SM30_RAT	SMP-30: Senescence marker protein-30	Internal sequence	721.71	5.11	830.10	34 051
135	SODC_RAT	Superoxide dismutase	AAA; confirmed by internal sequence (R. M. Van Frank)	1161.24	5.74	1388.68	18 173
172	TPM-RAT ^{e)}	Tm: tropomyosin	Location in cytoskeleton, 2-D position relative to human, Ab	476.24	4.66	957.86	28 865
277, 56	TBA1_RAT	Tubulin α	Positional homology with human through coelectrophoresis, cytoskeletal location	688.22	5.06	537.67	54 620
50, 1225	TBB1_RAT	Tubulin β	Positional homology with human through coelectrophoresis, cytoskeletal location	621.29	4.93	535.48	54 855
1224	VIME_RAT	Vimentin	Positional homology with human through coelectrophoresis, cytoskeletal location	673.00	5.03	539.50	54 426

Table 1. continued

MSN ^{a)}	Protein IDb)	Protein name	Identification comments	Gel X ^{c)}	Experimental pI ^{d)}	Gel Y ^{c)}	Experimental M _r ^{d)}
113	Unknown	?; not in sequence databases	Internal sequence	1191.28	5.78	680.42	42 469
104	BBPL_RAT	23 kDa morphine-binding protein	Internal sequence	773.31	5.20	1182.41	22 363

a) Master spot number (MSN) from [1]

b) SwissPROT identifier

c) Coordinates of the most basic or most abundant assigned spot on the F344 master gel pattern

d) pI and M_r of the most basic or most abundant assigned spot, derived from the calibration functions included here

e) SwissPROT style proposed identifier

Abbreviations: AAA, amino acid analysis; Ab, antibody

Table 2. Equations and coefficients

Function	Equation (f)	r ²	a	b	c	d	e
Rat gel Y = f(computed M _r)	$y = a + b \exp(-x/c)$	0.988181021	178.74803	1967.7892	32363.958		
Rat gel X = f(computed pI)	$y = a + bx + cx/\ln x + d/x + e/x^{1.5}$	0.99247216	-8685665.5	-904497.94	3856926.1	18276844	-27154534
Computed M _r = f(rat gel Y)	$y = a + bxc$	0.9960177	-8464.5809	19095881	-0.9086255		
Computed pI = f(rat gel X)	$y = a + bx + cx^2 + dx^2 \ln x + ex^3$	0.99176499	4.044686	-0.00114238	0.0000323	-0.00000455	0.00000000176
Mouse gel Y = f(rat gel Y)	$y = a + bx + cx^{1.5} + dx^{0.5} \ln x + ex/\ln x$	0.99951069	11861.44	678.91666	-0.78964914	1567.5639	-6953.9592
Mouse gel X = f(rat gel X)	$y = a + bx^2 \ln x + cx^{2.5} + dx^3$	0.99926349	58.935923	0.00091353	-0.000213688	0.00000159	
Rat gel Y = f(mouse gel Y)	$y = a + bx^2 \ln x + cx^{2.5} + dx^3$	0.99950032	69.740526	0.00050772	-0.000130392	0.00000116	
Rat gel X = f(mouse gel X)	$y = a + bx + cx^2 \ln x + dx^{2.5} + ex^3$	0.9992832	-198.07189	2.0899063	-0.000671191	0.000145189	-0.000000986

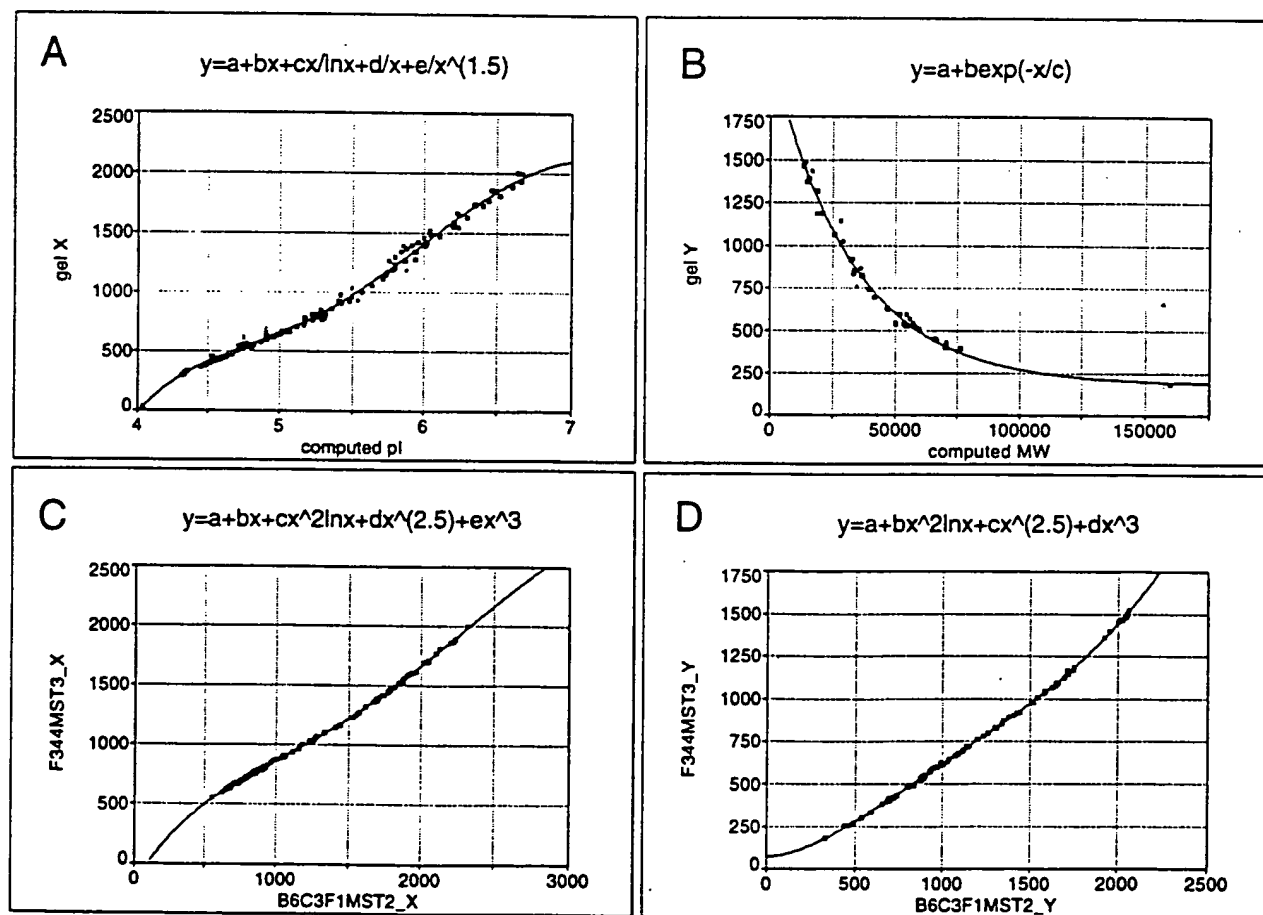


Figure 2. Plots showing fits of selected equations (continuous curves) to data on identified proteins (square symbols). (A) pI computed from sequence data versus gel X position for identified spots in F344 rat liver; (B) M_r computed from sequence data versus gel Y position for identified spots in F344 rat liver; (C) gel X position for spots in B6C3F1 mouse liver versus X position in F344 rat liver, for coelectrophoresing spots; (D) gel Y position for spots in B6C3F1 mouse liver versus Y position in F344 rat liver, for coelectrophoresing spots. In each case, inverse equations were also computed (Table 2).

B6C3F1 MOUSE LIVER 2-D PROTEIN PATTERN

v1.1 (B6C3F1MST2 Liver) 28-Apr-1995 © by Large Scale Biology Corporation,
9620 Medical Center Drive, Rockville, MD 20850 USA 301/424-5989
MW and computed pI scales derived from data to known proteins

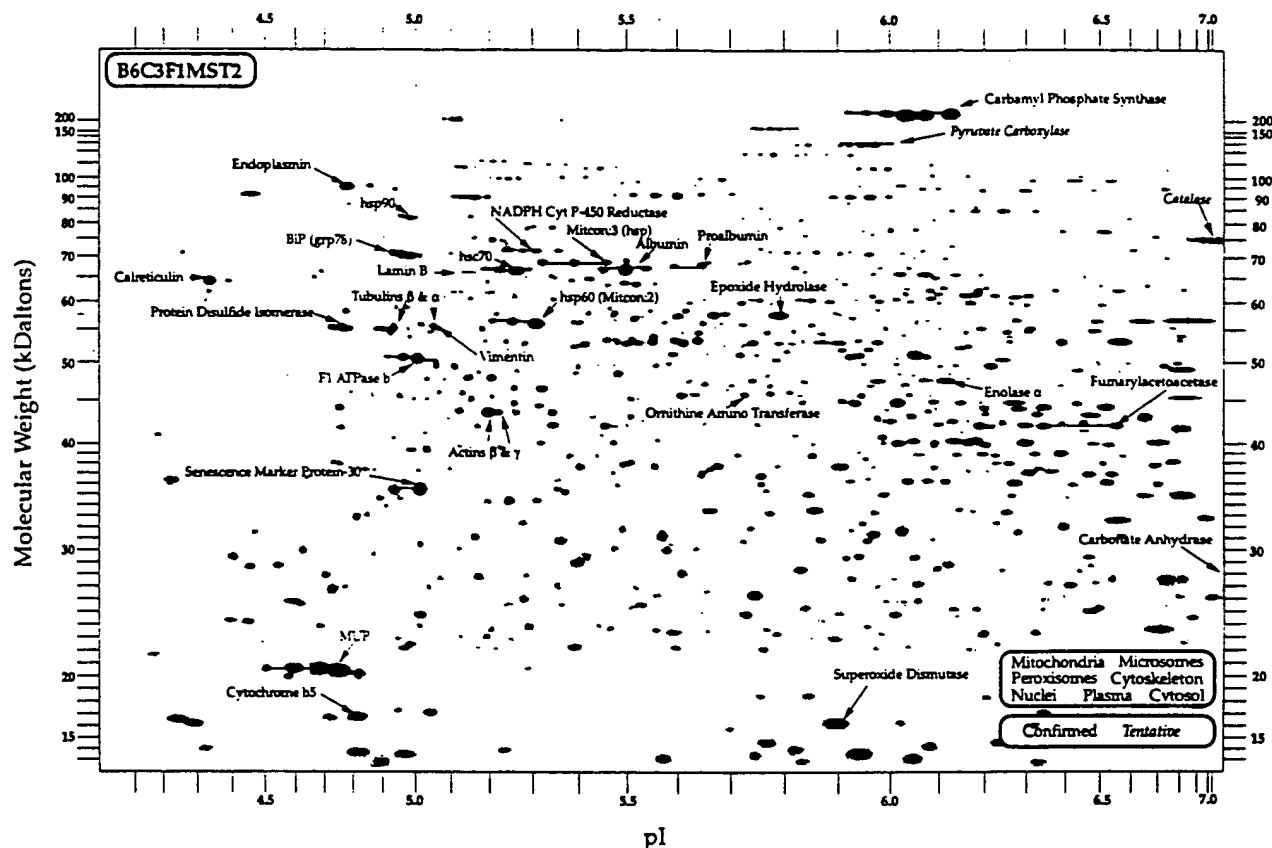


Figure 3. Master 2-D gel pattern for B6C3F1 mouse liver, standardized using the F344 rat liver pattern identifications, according to the method described in the text. Twenty-nine proteins are identified.

$$pI_{\text{RAT PLASMA}} = \frac{f_{\text{RAT LIVER}} \times pI_{\text{RAT PLASMA}} + \text{LIVER} \times \text{RAT LIVER} \times pI_{\text{RAT PLASMA}}}{f_{\text{RAT PLASMA}} \times \text{RAT PLASMA} + \text{LIVER} \times \text{RAT PLASMA}} \quad (8)$$

This unified approach, in which one well-populated 2-D pattern is used to standardize a family of other patterns, has the additional advantage that the resulting pI and M_r scales are directly compatible. Hence one can compare the relative pI 's of mouse and rat versions of a sequenced protein in a consistent pI measurement system, and select likely inter-species analogs based on positional relationships on common scales. Adoption of immobilized pH gradient (IPG) technology [4-7] will result in substantial improvements in pI positional reproducibility for standard 2-D maps such as those presented here; however, we believe that our approach will continue to be useful in establishing the empirical pH gradient actually achieved by such gels under given experimental conditions (temperature, urea concentration, etc.), in relating patterns run on different IPG ranges and using different lots of IPG gels (between which some variation will persist). Development of rodent organ maps is a continuing effort in our laboratories [8-10], and results in regular additions of identified proteins. Those who wish to receive current rodent liver maps, with color annotations, should send a stamped self-addressed envelope to the first author.

We would like to thank the individuals who provided antibodies mentioned in Table 1, and R. M. van Frank for unpublished sequenced data.

Received May 31, 1995

References

- [1] Anderson, N. L., Esquer-Blasco, R., Hofmann, J.-P., Anderson, N. G., *Electrophoresis* 1991, 12, 907-930.
- [2] Rosenfeld, J., Capdevielle, J., Guillemot, J. C., Ferrara, P., *Anal. Biochem.* 1992, 203, 173-179.
- [3] Witzmann, F., Clack, J., Fultz, C., Jarnot, B., *Electrophoresis* 1995, 16, 451-459.
- [4] Rosengren, A. E., Bjellqvist, B., Gasparic, V., *US Patent* 4130470, December 1978.
- [5] Gianazza, E., Artoni, G., Righetti, P. G., *Electrophoresis* 1983, 4, 321-326.
- [6] Görg, A., Postel, W., Günther, S., Weser, J., *Electrophoresis* 1985, 6, 599-604.
- [7] Gianazza, E., Astrua-Testori, S., Giacon, P., Righetti, P. G., *Electrophoresis* 1985, 6, 332-339.
- [8] Myers, T. G., Dietz, E. C., Anderson, N. L., Khairallah, E. A., Cohen, S. D., Nelson, S. D., *Chem. Res. Toxicol.* 1995, 8, 403-413.
- [9] Cunningham, M. L., Pippin, L. L., Anderson, N. L., Wenk, M. L., *Toxicol. Appl. Pharmacol.* 1995, 131, 216-223.
- [10] Anderson, N. L., Copple, D. C., Bendele, R. A., Probst, G. S., Richardson, F. C., *Fundam. Appl. Toxicol.* 1992, 18, 570-580.